

Seifert, Andreas; Hilligus, Annegret Helen; Schaper, Niclas  
**Entwicklung und psychometrische Überprüfung eines Messinstruments  
zur Erfassung pädagogischer Kompetenzen in der universitären  
Lehrerbildung**

*Lehrerbildung auf dem Prüfstand 2 (2009) 1, S. 82-103*



Quellenangabe/ Reference:

Seifert, Andreas; Hilligus, Annegret Helen; Schaper, Niclas: Entwicklung und psychometrische Überprüfung eines Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung - In: Lehrerbildung auf dem Prüfstand 2 (2009) 1, S. 82-103 - URN: urn:nbn:de:0111-pedocs-147005 - DOI: 10.25656/01:14700

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-147005>

<https://doi.org/10.25656/01:14700>

in Kooperation mit / in cooperation with:



[www.vep-landau.de](http://www.vep-landau.de)

#### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

#### Kontakt / Contact:

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft



**Herausgeber**

Rainer Bodensohn, Reinhold S. Jäger und Andreas Frey  
Bürgerstraße 23, 76829 Landau/Pfalz  
Telefon: +49 6341 906 165, Telefax: +49 6341 906 166

**Verlag**

Empirische Pädagogik e. V.  
Bürgerstraße 23, 76829 Landau/Pfalz  
Telefon: +49 6341 906 180, Telefax: +49 6341 906 166  
E-Mail: [info@vep-landau.de](mailto:info@vep-landau.de)  
Homepage: <http://www.vep-landau.de>

**Umschlaggestaltung**

© Harald Baron

**Druck**

DIFO Bamberg

*Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, werden vorbehalten. Kein Teil des Werks darf in irgendeiner Form (durch Fotografie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung des Verlags reproduziert oder unter Verwendung elektronischer Systeme verbreitet werden.*

ISSN 1867-2779

ISBN 978-3-941320-11-6

© Verlag Empirische Pädagogik, Landau 2009

---

---

## Inhalt

---

---

### Editorial

- Schaper, N., Hilligus, A. H. & Reinhold, P.:  
Kompetenzmodellierung und -messung in der Lehrerbildung ..... 1

### Originalarbeiten

- Schott, F & Azizi Ghanbari, S.:  
Modellierung, Vermittlung und Diagnostik der Kompetenz kompetenzorientiert zu unterrichten – wissenschaftliche Herausforderung und ein praktischer Lösungsversuch ..... 10
- Heinzer, S., Oser, F. & Salzmann, P.:  
Zur Genese von Kompetenzprofilen ..... 28
- Schmelzing, S., Fuchs, C., Wüsten, S., Sandmann, A. & Neuhaus, B.:  
Entwicklung und Evaluation eines Instruments zur Erfassung des fachdidaktischen Reflexionswissens von Biologielehrkräften ..... 57
- Seifert, A., Hilligus, A. H. & Schaper, N.:  
Entwicklung und psychometrische Überprüfung eines Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung ..... 82
- Riese, J. & Reinhold, P.:  
Fachbezogene Kompetenzmessung und Kompetenzentwicklung bei Lehramtsstudierenden der Physik im Vergleich verschiedener Studiengänge ..... 104
- König, J. & Blömeke, S.:  
Disziplin- oder Berufsorientierung?  
Zur Struktur des pädagogischen Wissens angehender Lehrkräfte ..... 126
- Schmotz, C. & Blömeke, S.:  
Zum Verhältnis von fachbezogenem Wissen und epistemologischen Überzeugungen bei angehenden Lehrkräften ..... 148
- Schaper, N.:  
Aufgabenfelder und Perspektiven bei der Kompetenzmodellierung und -messung in der Lehrerbildung ..... 166

- Impressum**..... 200

---

---

## Contents

---

---

### Articles

- Schott, F & Azizi Ghanbari, S.:  
Modeling, imparting and diagnosing the competence to teach  
competence-oriented – scientific challenge and a practical approach ..... 10
- Heinzer, S., Oser, F. & Salzmann, P.:  
On the genesis of competency profiles of teachers in the field of  
vocational education and training ..... 28
- Schmelzing, S., Fuchs, C., Wüsten, S., Sandmann, A. & Neuhaus, B.:  
Development and evaluation of a test to measure pedagogical content  
knowledge of biology teachers ..... 57
- Seifert, A., Hilligus, A. H. & Schaper, N.:  
Development and psychometrical testing of an instrument for the use  
of measuring pedagogical competence in university teacher education .... 82
- Riese, J. & Reinhold, P.:  
Measurement of future physics teachers' professional knowledge  
and its development within different teacher education programs ..... 104
- König, J. & Blömeke, S.:  
Orientation towards discipline or profession?  
On the structure of future teachers' pedagogical knowledge ..... 126
- Schmoltz, C. & Blömeke, S.:  
On the relationship of future teachers' content knowledge,  
pedagogical content knowledge and epistemological beliefs ..... 148
- Schaper, N.:  
Scope and perspectives of competence modelling and measurement  
in teacher education research ..... 166

---

## Impressum

---

**Lehrerbildung auf dem Prüfstand**

ISSN 1867-2779

ISBN 978-3-941320-11-6

Bürgerstraße 23, 76829 Landau/Pfalz

Telefon: +49 6341 906 180 Telefax: +49 6341 906 166

E-Mail: lbp@vep-landau.de Homepage: <http://www.vep-landau.de>**Erscheinungsweise/Preis**

Die Zeitschrift erscheint zweimal jährlich. Der Bezugspreis beträgt € 36,00/Jahr (Studierende: € 18,00/Jahr) zzgl. Porto. Kündigung 6 Wochen vor Jahresende.

**Herausgeber**

Rainer Bodensohn (Landau), Andreas Frey (Mannheim), Reinhold S. Jäger (Landau)

**Beirat**

Karl-Heinz Arnold (Hildesheim), Franz Baeriswyl (Fribourg), Günter Dörr (St. Ingbert), Ludwig Haag (Bayreuth), Niclas Schaper (Paderborn)

**Redaktion**

Ines Weresch-Deperrois

**Beiträge**

Die Zeitschrift *Lehrerbildung auf dem Prüfstand* (LbP) versteht sich als Organ, das die Lehrerbildung im gesamten deutschsprachigen Raum mit Hilfe empirischer Untersuchungen, kritischer Diskurse und Diskussionen begleitet und evaluiert. Das Ziel besteht darin, der Lehrerbildung zu einem hohen Niveau zu verhelfen. Dieses Ziel wird sowohl mit Themenheften verfolgt, die sich einer bestimmten Thematik widmen, als auch mit Heften, in denen empirische Untersuchungen, historische Diskurse, Perspektiven der Lehrerbildung, Metaanalysen, Buchbesprechungen etc. abgedruckt werden.

Die Zeitschrift veröffentlicht Beiträge generell zu Inhalten der Lehrerbildung. Die Beiträge sind im Regelfall empirisch orientiert, Beiträge mit grundsätzlichem Charakter sind erwünscht.

**Manuskriptgestaltung**

Die aktuellen Autorenrichtlinien finden Sie im Internet unter <http://www.vep-landau.de>. Im Übrigen müssen die eingereichten Beiträge den „Richtlinien zur Manuskriptgestaltung“ (Göttingen: Hogrefe, 2007) entsprechen. Abgabe des Manuskripts in dreifacher Ausfertigung oder per E-Mail als Word- oder pdf-Datei. Endfassung (Text und Abbildungen) auf CD-Rom oder als E-Mail-Anhang.

---

## Originalarbeiten

---

Andreas Seifert, Annegret Helen Hilligus und Niclas Schaper

### Entwicklung und psychometrische Überprüfung eines Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung

**Zusammenfassung:** Dieser Beitrag dokumentiert die theoriegeleitete Entwicklung eines objektiven und ökonomischen Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung. Pädagogische Kompetenz wird neben Fachwissen, fachdidaktischem Wissen, Einstellungen und motivationalen Aspekten als Teil eines umfassenderen Konzepts professioneller Lehrerkompetenz angesehen. Pädagogische Kompetenz lässt sich in verschiedene Inhalts- sowie Anforderungsbereiche aufteilen. Diese Struktur ist die Basis für die Operationalisierung der pädagogischen Kompetenz beim vorliegenden Instrument. Auf der Grundlage einer Stichprobe von 700 angehenden Lehrerinnen und Lehrern wird das Instrument bezüglich seiner itemanalytischen Kennwerte eingeschätzt. Weiterhin wird eine kriteriengeleitete Kürzung des Instruments vorgenommen, um die pädagogischen Kompetenzen im Zusammenhang mit den anderen Bestandteilen professioneller Lehrerkompetenz erheben zu können. Weitergehende Fragestellungen und Einsatzgebiete der Langform sowie der Kurzform werden diskutiert.

**Schlagwörter:** Lehrerausbildung – Leistungstest – pädagogische Kompetenz

### Development and psychometrical testing of an instrument for the use of measuring pedagogical competence in university teacher education

**Summary:** This report documents the theory-based development of an objective and economical instrument for the use of measuring pedagogical competence in university teacher education. Pedagogical competence is viewed as part of a broader concept of professional competence of teachers besides content knowledge, pedagogical content knowledge, professional attitudes and motivational aspects. Pedagogical competence can be divided in several domains of content and requirement. This structure is the foundation of the operationalization of pedagogical competence for the instrument at hand. On the basis of a sample of 700 future teachers the instrument is assessed regarding item-analytic values. Furthermore a criterion-related reduction is conducted. With this short version pedagogical knowledge can be measured in connection with the other components of professional competence of teachers. Advanced questions and fields of application are discussed for the long version and the short version.

**Key words:** achievement test – pedagogical competence – teacher education

## 1. Hintergrund

In Deutschland erfährt die Lehrerausbildung in den letzten Jahren durch die übergreifenden Initiativen der Kultusministerkonferenz (KMK) (z. B. Terhart, 2002; KMK 2004, 2008) eine konzeptionelle, inhaltliche und strukturelle Neuorientierung. Es findet ein Umdenken hin zu einer *output*-Steuerung statt, die sich in einer zunehmenden Orientierung der Curricula an Standards und Kompetenzen manifestiert (vgl. Hilligus & Rinkens, 2006). Die KMK hat in diesem Feld eine

Rahmen gebende Aufgabe übernommen. In der Erziehungswissenschaft hat sie bereits 2004 die „Standards für die Lehrerbildung: Bildungswissenschaften“ verabschiedet. Für die Fächer sind 2008 „ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung“ verabschiedet worden. Diese Rahmenvorgaben werden über die Bundesländer hinweg in einer neuen Weise normierend wirken, da sie zukünftig als Grundlage für die Akkreditierung von Lehramtsstudiengängen dienen sollen.

Mit dem Wandel in der Ausbildung einher geht eine Umorientierung in der Lehrerbildungsforschung. Angeregt durch andere Bereiche der Bildungsforschung wie etwa die internationalen Schulleistungsstudien PISA und TIMSS, haben die Anstrengungen, Kompetenzen mit standardisierten Testverfahren zu überprüfen, auch in der Lehrerausbildung merklich zugenommen. Im Vordergrund steht dabei derzeit einerseits die Entwicklung von Messinstrumenten, um die empirische Wirksamkeit der Lehrerausbildung zu überprüfen (vgl. Blömeke, 2004). Andererseits bedarf es entsprechender Messinstrumente, um in Anlehnung an die internationale Diskussion auch in der deutschen Lehrerausbildung Ansätze zur *output*-Steuerung aufzubauen (z. B. Darling-Hammond & Bransford, 2005).

An der Universität Paderborn hat eine Neuorientierung der Lehrerbildung in Orientierung an Kompetenzen und Standards im Rahmen des Projekts SPEE (Standards – Profile – Entwicklung – Evaluation) stattgefunden, das vom Zentrum für Bildungsforschung und Lehrerbildung der Universität Paderborn (PLAZ) konzipiert wurde (Hilligus, Reinhold, Rinkens & Tulodziecki, 2004). Parallel zur Entwicklungsarbeit entstand das Interesse, Hinweise über den Kompetenzerwerb in der ersten Phase der Lehrerbildung zu erhalten. Dazu werden in einem interdisziplinär ausgerichteten Forschungskolleg „Kompetenzentwicklung und Messung“ (KEM) Instrumente zur Messung professioneller Handlungskompetenz bei Studierenden des Lehramts entwickelt. Da die entwickelten Instrumente über Paderborn hinaus einsetzbar sein sollen, wird bei der Konzeption des Kompetenzrahmenmodells und bei der Formulierung der Iteminhalte auf die Übertragbarkeit auf andere Ausbildungskontexte geachtet. Neben Instrumenten zur Erfassung von Kompetenzen in den Teilbereichen Fachwissen und Fachdidaktik für die Fächer Mathematik, Physik, Deutsch, Katholische Religionslehre und Sport wird ein Verfahren konzipiert, das in der Lage ist, den Kompetenzerwerb im erziehungswissenschaftlichen Studium mithilfe eines eigenständigen Tests zu ermitteln.

## 2. Zielsetzung der Instrumentenentwicklung

Der vorliegende Beitrag beschäftigt sich mit der Entwicklung eines Messinstrumentes zur Erfassung pädagogischer Kompetenzen, welches in der Lage ist, die Kompetenzentwicklung bzw. den Kompetenzerwerb bezüglich des erziehungswissenschaftlichen Teils des Studiums abzubilden und damit auch als Eva-



luationsinstrument im Rahmen dieses (Teil-)Studiengangs eingesetzt zu werden. Das mit dem Vorhaben verbundene Interesse besteht darin, über die bei der Kompetenzevaluation derzeit noch dominierenden Selbsteinschätzungsverfahren hinauszugehen und sie durch Messansätze zu ergänzen, um verlässlichere Ergebnisse über Kompetenzausprägungen und -veränderungen in universitären Bildungsprozessen zu erhalten. Im Zentrum steht dabei die Frage, inwieweit die Lehrerausbildung wirksam ist: Werden im Studium pädagogische Kompetenzen erworben, und lässt sich im Verlaufe des Studiums ein Kompetenzzuwachs erkennen? Neben der validen Erfassung von Kompetenzen und ihres Erwerbs im erziehungswissenschaftlichen Lehramtsstudium an der Universität Paderborn soll das zu entwickelnde Instrument auch zur Kompetenzerfassung und Evaluation der Ausbildung an anderen Standorten geeignet sein. Die Übertragbarkeit des Messansatzes wird durch gesonderte empirische Studien überprüft.

Weiterhin soll auf der Basis einer Langform des Instruments zur Messung pädagogischer Kompetenzen eine Kurzform erstellt werden, um in Verbindung mit den oben genannten fachspezifischen Kompetenzmessinstrumenten sowie weiteren Fragebögen zu professionsbezogenen Überzeugungen und motivationalen Orientierungen eine Erhebung und Diagnose umfassender Kompetenzprofile für das Lehramtsstudium zu ermöglichen. Mithilfe solcher umfassenden Kompetenzerhebungen können in einem weiteren Schritt auch die Zusammenhänge pädagogischer Kompetenzen zu den anderen Komponenten professioneller Lehrerkompetenz überprüft werden.

### 3. Grundlegendes Kompetenzstrukturmodell

Die der Instrumentenentwicklung zugrunde liegende Kompetenzauffassung und das entsprechende Kompetenzstrukturmodell (siehe Abbildung 1) basieren auf aktuellen Ansätzen der empirischen Bildungsforschung (Bromme, 1997; Klie-me et al, 2003; Shulman, 1986; Weinert, 2001). Auf entsprechende begriffliche Bestimmungen und Modelle beziehen sich auch die großen Lehrerbildungsstudien MT21 und COACTIV, deren Autoren vergleichbare Ansätze und Strukturmodelle entwickelt haben, die ihnen als heuristische Rahmenmodelle zur Kompetenzanalyse dienen (vgl. z. B. Baumert & Kunter, 2006).

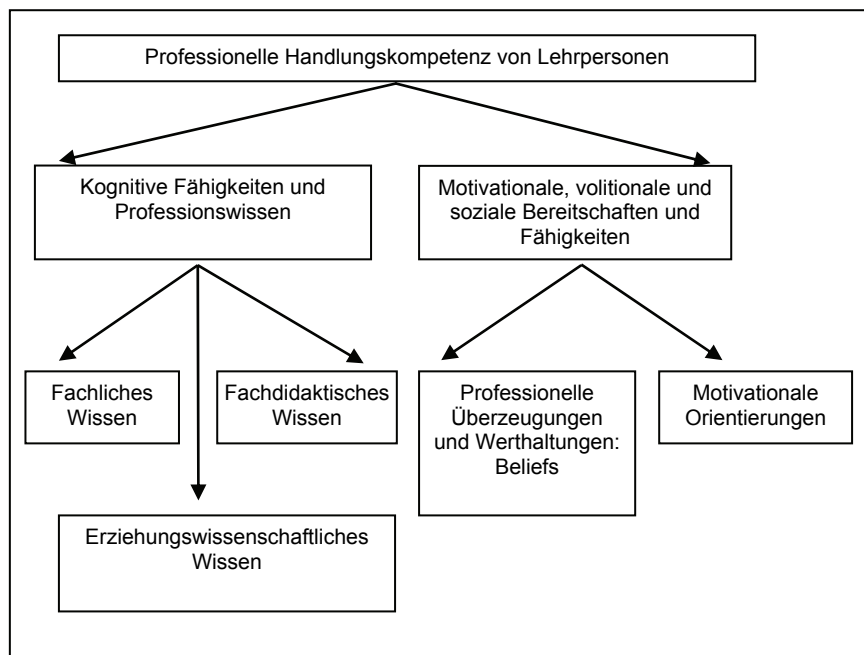


Abbildung 1: Grundlegendes Kompetenzstrukturmodell zur Instrumentenentwicklung

Kompetenzen können nach Weinert (2001, 27f.) verstanden werden als „die bei Individuen verfügbaren oder durch sie erlernbaren Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“

Es wird somit davon ausgegangen, dass sich Kompetenzen bei Lehrpersonen aus wissensbasierten Anteilen sowie aus komplementären Bereitschaften und Einstellungen zusammensetzen (Beliefs und motivationale Orientierungen).

Die Strukturierung im Bereich der kognitiven Fähigkeiten in fachliches, fachdidaktisches und erziehungswissenschaftliches Wissen geht auf die Topologie von Wissensdomänen im Lehrerberuf nach Shulman (1986, 1985) zurück, der „content knowledge“, „pedagogical content knowledge“ und „pedagogical knowledge“ unterscheidet. Diese Einteilung entspricht damit auch der curricularen Strukturierung von Lehramtsstudiengängen in fachliche, fachdidaktische und erziehungswissenschaftliche Bereiche bzw. Module, wie sie in Deutschland üblich

ist. Auf dieser Grundlage kann die Messung professioneller Handlungskompetenz oftmals zugleich auch als Evaluation eines Lehrer ausbildenden Curriculums bzw. bestimmter Teile eines entsprechenden Curriculums angelegt werden.

#### 4. Grundlagen zur Operationalisierung von pädagogischer Kompetenz

Für die Erfassung fachlicher und fachdidaktischer Kompetenzen bei angehenden Lehrkräften liegen in Teilbereichen bereits verschiedene Operationalisierungsansätze und Instrumente vor, die sich nicht nur auf Selbsteinschätzungen der Studierenden beziehen. Dies gilt insbesondere für das Fach Mathematik (siehe insbesondere Brunner et al., 2006 oder Blömeke et al., 2008a). Vergleichbare Instrumentarien für den Bereich der pädagogischen Kompetenz liegen jedoch nur in ersten Ansätzen vor.

Einerseits existieren hierzu eine Reihe von theoriegeleiteten Vorschlägen zur Operationalisierung, die insbesondere vor dem Hintergrund entstanden sind, Standards für die erziehungswissenschaftliche Ausbildung von Lehrkräften zu formulieren (siehe z. B. Baumert & Kunter, 2006; Tulodziecki, 2004; Oser & Oelkers, 2001). Auf dieser Grundlage wurden zur Erfassung von Standards Fragebögen entwickelt, die Selbst- oder Fremdeinschätzungen zur Ausprägung solcher Standards bzw. Kompetenzen bei Lehrkräften in unterschiedlichen Ausbildungsphasen erfragen (z. B. Abs, Döbrich, Vögele & Klieme, 2005; Baer et al., 2007; Frey, 2008; Schubarth & Polenz, 2006; Oser & Oelkers, 2001). Diese Studien haben erste Hinweise im Hinblick auf den Erwerb pädagogischer Kompetenz gegeben. Die Validität solcher Self-Assessment-Verfahren wird in Bezug auf die Erfassung von Kompetenzen allerdings eher kritisch diskutiert und es wird gefordert, sie durch objektive Leistungstests zu ersetzen bzw. zu ergänzen (vgl. Schaefer, 2002; Terhart, 2002).

Operationalisierungsansätze und Instrumente, die pädagogische Kompetenz mithilfe von objektiven Leistungstests versuchen zu erfassen, sind jedoch kaum vorhanden oder befinden sich parallel zu dem hier dargestellten Ansatz in der Entwicklung (vgl. König & Blömeke, im Druck und in diesem Band; Dubberke, Kunter & McElvany, 2008 oder Blömeke, Felbrich & Müller, 2008). Wie dieser Bereich messmethodisch operationalisiert werden kann und wie er strukturiert ist, kann vor dem Hintergrund vorhandener Studien noch nicht schlüssig beantwortet werden (vgl. Blömeke, Felbrich & Müller, 2008). Im Rahmen der MT 21 Studie gliedern Blömeke et al. (2008a) auf der Grundlage einer Analyse curricularer Strukturen im Fach Erziehungswissenschaft pädagogische Kompetenz in drei Teilbereiche: allgemeinpädagogische, pädagogisch-psychologische und bildungssoziologische Lerninhalte und Fragestellungen. Die Beschränkung auf diese Teilbereiche wird mit der begrenzten Testzeit und der Vielfalt der sechs Teilnehmer-

länder dieser internationalen Studie begründet, die es nötig machte, sich auf die in allen Ländern vorhandenen Kerngebiete zu beschränken, wohl wissend, dass dadurch das deutsche Spektrum erziehungswissenschaftlicher Curricula nicht vollständig erfasst wird. Diese drei Bereiche werden mithilfe von wissensbezogenen Testitems abgefragt, die sowohl geschlossene Formate in Form von Multiple-Choice-Items als auch offene Formate in Form von schriftlich zu beantwortenden Fragen zu unterrichtsbezogenen Anforderungssituationen beinhalten. Im Rahmen einer Weiterentwicklung dieses Ansatzes haben König und Blömeke (im Druck und in diesem Band) ein umfangreicheres Instrument zur Messung der pädagogischen Kompetenz entwickelt, das in Anlehnung an Baumert et al. (2004), Gruehn (2000) und Helmke (2003) übergreifende pädagogische Anforderungen von Lehrkräften in fünf Anforderungs- bzw. Inhaltsbereiche gliedert: Strukturierung von Unterricht, Motivierung, Umgang mit Heterogenität, Klassenführung und Leistungsbeurteilung. Die insgesamt 50 Testaufgaben beinhalten sowohl geschlossene als auch offene Item-Formate, um deklaratives und prozedurales Wissen für diese Anforderungs- bzw. Inhaltsdomänen zu erfassen. Weiterhin wird zurzeit ein Instrumentarium zur Messung pädagogischer Kompetenz im Rahmen der MPI Arbeitsgruppe um Baumert und Kunter entwickelt (Dubberke, Kunter & McElvany, 2008). Pädagogische Kompetenz wird dabei ebenfalls in die fünf Inhaltsbereiche gemäß Baumert et al. (2004) bzw. König und Blömeke (im Druck) eingeteilt. Neben wissensbezogenen Items werden in diesem Zusammenhang allerdings auch Videovignetten zur Operationalisierung von Beurteilungs- und Handlungsanforderungen in authentischen Unterrichtssituationen und zur Erfassung entsprechender Leistungen eingesetzt.

Als Grundlage für die Itementwicklung zur Messung pädagogischer Kompetenz von zukünftigen Lehrpersonen im Kontext des SPEE-Projekts wurde ein fachinhaltlich- und dimensionsspezifisch ausgestaltetes Rahmenmodell entworfen. Es dient als heuristisches Instrument zur Itementwicklung, um sicherzustellen, dass zentrale kompetenzrelevante Inhalte des Studiums durch die Items repräsentiert werden. Das Rahmenmodell wurde vor dem Hintergrund folgender Leitfragen konzipiert:

- Welche erziehungswissenschaftlichen Inhaltsbereiche umfasst die Ausbildung von zukünftigen Lehrpersonen gemäß den Rahmenvorgaben der KMK, des Landes NRW und dem modularisierten Paderborner Curriculum?
- Welche inhaltlichen Strukturierungen des erziehungswissenschaftlichen Wissens sind dabei zu berücksichtigen?
- Mit welchen unterrichts- und diagnosebezogenen Aufgaben werden Lehrpersonen konfrontiert?

- Welche kognitiven Aktivitäten bzw. Anforderungen in Anlehnung an Terhart (2002) können beim Umgang mit diesen Inhaltsbereichen unterschieden werden? Das resultierende Rahmenmodell ist in Tabelle 1 wiedergegeben.

Tabelle 1: Rahmenmodell zur Strukturierung pädagogischer Kompetenzen mit stichwortartig formulierten Beispielen der zugehörigen pädagogischen Kompetenzen in den jeweiligen Zellen der Matrix

<b>Kognitive Anforderungen</b> <b>Inhaltsbereiche</b>	<b>Wissen reproduzieren</b> <b>Verstehen</b>	<b>Reflektieren</b> <b>Anwenden</b>	<b>Urteilen</b> <b>Bewerten</b> <b>Entscheiden</b>
<b>Erziehung und Bildung</b>	pädagogische Strömungen kennen Aussagen in Kommunikationsmodell einordnen	Aktivitäten Handlungstheorien zuordnen Lernmechanismen Fällen zuordnen	mithilfe psychologischer Lerntheorien Handlungsalternativen generieren
<b>Unterricht und Allgemeine Didaktik</b>	Aufgabentypen kennen Merkmale guten Unterrichts nennen	Unterrichtsphasen Situationen zuordnen Motivationsarten bezüglich Lernförderlichkeit ordnen	Strukturierung von Unterricht modellieren Umgang mit Fehlern in Situation darstellen
<b>Schulentwicklung und Gesellschaft</b>	zentrale Aufgaben im Lehrerberuf kennen Ziele von Schulentwicklung nennen	Leistungsnormen Zielen zuordnen Problematische Situationen in Bezug auf Gütekriterien erkennen	Argumente für externe und interne Differenzierung beurteilen Auswertungsobjektivität bewerten

In den Zeilen des Rahmenmodells befinden sich die Inhaltsbereiche, wie sie im Paderborner Curriculum konzeptualisiert sind (vgl. Tulodziecki, 2004). Inhaltlich sind alle Kompetenzbereiche abgedeckt, die in den KMK-Standards (2004) genannt werden: Unterrichten, Erziehen, Beurteilen und Beraten, Innovieren. Das Unterrichten ist in dieser Matrix Gegenstand von Unterricht und Allgemeiner Didaktik, Erziehen gehört zum Bereich Bildung und Erziehung. Beurteilen und Beraten sowie Innovieren sind dem Bereich Schulentwicklung und Gesellschaft zuzuordnen, wobei das Beurteilen und Beraten zusätzlich auch Gegenstand von Unterricht und Allgemeiner Didaktik ist. Um den pädagogischen Horizont des Berufsfeldes „Lehrerin/Lehrer“ möglichst umfassend abzubilden, wird das Konstrukt „Pädagogische Kompetenz“ in diesem Ansatz somit weiter gefasst als bei den oben genannten Instrumentenentwicklungen. Die im Ansatz von König und Blömeke (im Druck) sowie Dubberke, Kunter und McElvany (2008) erfassten

fünf Inhaltsbereiche pädagogischer Kompetenz werden im vorliegenden Rahmenmodell in erster Linie im Bereich „Unterricht und Allgemeine Didaktik“ erfasst. Die bei Blömeke et al. (2008) berücksichtigten drei Inhaltsbereiche beziehen sich hinsichtlich der bildungssoziologischen Facette zusätzlich auch auf den Bereich „Erziehung und Bildung“ bzw. „Schulentwicklung und Gesellschaft“ des beschriebenen Rahmenmodells.

Die Spalten des Rahmenmodells geben die verschiedenen kognitiven Aktivitäten bzw. Anforderungen an die Nutzung des professionellen Wissens in Anlehnung an die Kategorien von Terhart (2002) wieder. Der Bereich der „Kommunikation“ nach Terhart (2002) wird aus Gründen der Handhabbarkeit vernachlässigt, da sich die Messung dieses Anforderungsbereichs nur schwer mithilfe eines standardisierten Paper-and-Pencil-Verfahrens realisieren lässt.

Zu dem Modell von Blömeke et al. (2008a) gibt es darüber hinaus Unterschiede in den strukturellen Dimensionen. Während Blömeke die drei Inhaltsbereiche an den Hauptdisziplinen Pädagogik, Psychologie und Soziologie ausrichtet, die das erziehungswissenschaftliche Studium verantworten, verzichtet das Paderborner Modell auf eine disziplinäre Zuordnung. Ersetzt wird diese durch eine thematische Klassifizierung in Erziehung und Bildung, Unterricht und Allgemeine Didaktik sowie Schulentwicklung und Gesellschaft, die curricular jeweils in Modulen abgebildet wird und die von den Disziplinen gemeinsam mit Angeboten bedient werden. Die in den Spalten der Paderborner Matrix abgetragenen drei kognitiven Anforderungsdimensionen werden bei Blömeke et al. (2008a) in den zwei Dimensionen des deklarativen und prozeduralen Wissens konzeptualisiert, die dort zur Aufgabenentwicklung herangezogen werden und in gewisser Weise die dritte Dimension darstellen. Im Paderborner Modell ist diese dritte Dimension durch inhaltliche Anforderungen bestimmt, die allerdings – gemäß den KMK-Vorgaben – weiter gefasst sind als in den anderen Modellen und nicht nur das Classroom Management umfassen.

Bezüglich der Zusammenhänge zwischen den Dimensionen wird Folgendes angenommen: Während für die Inhaltsdimensionen postuliert wird, dass sie untereinander eher als unabhängig anzusehen sind, wird bei den Anforderungsstufen eine hierarchische Struktur angenommen in dem Sinne, dass die Stufe „Wissen reproduzieren, Verstehen“ als Grundlage für die Stufe „Reflektieren, Anwenden“ angesehen werden kann, welche wiederum die Grundlage für die Stufe „Urteilen, Bewerten, Entscheiden“ darstellt. Außerdem wird angenommen, dass Inhalts- und Anforderungsdimensionen sich zueinander unabhängig verhalten. In den sich aus der Kombination von Inhalts- und Anforderungsbereichen ergebenden Zellen des Rahmenmodells (siehe Tabelle 1) werden in exemplarischer Form Iteminhalte bzw. durch die Items abgefragte Kompetenzen genannt. Die Iteminhalte beziehen

sich vornehmlich auf Lerninhalte der universitären Phase der Lehrerausbildung, für die das Instrument in einem ersten Schritt konstruiert wurde.

## 5. Psychometrische Messung erziehungswissenschaftlicher Kompetenz

Die Diagnostik bzw. Erfassung von Kompetenzen mithilfe von Leistungstests stellt besondere Ansprüche an die verwendeten Testverfahren. Zur Messung der kognitiven Facetten von Kompetenzen ist daher zu beachten, dass die entsprechenden Testformate nicht nur Anforderungen an die Reproduktion oder Rekonstruktion kompetenzrelevanter Wissenskonzepte und mentaler Operationen widerspiegeln. Von entscheidender Bedeutung ist bei der Kompetenzmessung auch, dass die Anwendung bzw. der Transfer von Wissen in die Praxis gefordert wird, d. h., dass auch Anforderungen an die erfolgreiche Lösung von Problemen in konkreten Situationen mithilfe von professionellem Wissen (Bromme, 1997) gestellt werden.

Dieses Verständnis des Konstrukts „Kompetenz“ stellt Testkonstrukteure vor schwierige Abwägungen bezüglich der Standardisierung und Validität der Testformate und -inhalte. Entsprechende Überlegungen in Bezug auf geeignete Testformate und -inhalte stehen meist im Spannungsfeld zwischen hoher Standardisierung aber eingeschränkter ökologischer Validität (z. B. bei Wissenstests mithilfe von Multiple-Choice-Items) und hoher ökologischer Validität aber eingeschränkter Standardisierung, wie an dem unten skizzierten Einsatz von Situational Judgement Tests deutlich wird.

Für die Messung von Kompetenzen gibt es unterschiedliche Möglichkeiten der Operationalisierung (Schaper, 2006). Unterschieden werden können Beobachtungsverfahren in natürlichen und künstlichen Settings (1), Methoden zur Selbst- und Fremdbeurteilung (2), biografische Ansätze (3), persönlichkeitsorientierte Ansätze (4) und aufgabenanalytische Ansätze (5).

Für eine quantitativ angelegte und große Stichproben umfassende Kompetenzmessung in der Lehrerausbildung sind persönlichkeitsorientierte (4) und aufgabenanalytische (5) Ansätze am ehesten geeignet: Persönlichkeitsorientierte Ansätze bieten sich für die Messung von Einstellungen/Überzeugungen und motivationalen Aspekten an. Aufgabenanalytische Ansätze können zur Messung von fachlichen, fachdidaktischen und pädagogischen Kompetenzkomponenten dienen.

Die anderen Verfahren kommen für Zwecke der Kompetenzmessung in der Lehrerbildung nur bedingt in Frage. Beobachtungsverfahren (1) sind durch einen hohen Durchführungs- und Auswertungsaufwand gekennzeichnet, so dass eine Testung von großen Fallzahlen auf dieser Basis kaum möglich ist. Methoden zur Selbst- und Fremdbeurteilung (2) erfüllen zwar durchaus Kriterien der Standardisierung, Handhabbarkeit und Ökonomie bei der Erfassung von Kompetenzen.

Allerdings sind solche Verfahren – wie oben bereits angesprochen – nur eingeschränkt valide, um präzise und belastbare Aussagen über Kompetenzausprägungen und damit verbundenen Leistungsdispositionen zu machen. Die Einschränkungen der Validität sind vor allem bedingt durch Beurteilungsfehler und Verzerrungstendenzen, denen Selbst- und Fremdeinschätzungen von Kompetenzen unterliegen (Nerdinger, Blickle & Schaper, 2008). Biografische Ansätze (3) eignen sich für eine standardisierte Erfassung von Kompetenzen ebenfalls nicht, da das dabei generierte Datenmaterial eher „qualitativer“ Natur ist und für eine quantitativ-orientierte vergleichende Auswertung nur bedingt verwendet werden kann.

Allerdings ist man bei der Erfassung pädagogischer Handlungskompetenz mit der Schwierigkeit konfrontiert, dass es nicht nur um die Abfrage von Wissen geht, sondern um Anforderungen, erfolgreiche Handlungspläne für bestimmte professionelle Anforderungssituationen zumindest entwickeln zu können. Daher eignet sich der aufgabenanalytische Ansatz insofern, als zwischen sog. Wissenstests und Situational Judgement Tests unterschieden werden kann. Wissenstests beruhen auf der Annahme, dass als zentrale Voraussetzung für kompetentes Handeln Kenntnisse über Handlungsbedingungen und Handlungsabläufe erforderlich sind. Auf dieser Grundlage wird somit von der Ausprägung des Wissens auf Kompetenzen geschlossen. Als Aufgabenformat werden zumeist Multiple-Choice-Items, Zuordnungsaufgaben, Kurzantworten oder Kurzesays verwendet. Mit Wissenstests wird in erster Linie Fakten- bzw. deklaratives Wissen abgefragt. Teilweise wird mithilfe solcher Testformate aber auch Zusammenhangs- und Handlungswissen erfasst.

Bei Situational Judgement Tests (Weekly & Ployhart, 2006) werden in standardisierter Form Anforderungssituationen bzw. hypothetische Szenarien vorgegeben, bei denen die Probanden gefordert sind, die (Unterrichts-)Situation zu analysieren und angemessene Verhaltensweisen zur Problemlösung zu generieren. Mithilfe solcher Szenarien wird daher von den Testteilnehmern verlangt, dass sie ihr Wissen situationsangemessen anwenden können. Die Antworten werden zumeist anhand von situationsbezogenen Kriterien und Musterlösungen bewertet (Ulbricht, Schaper & Hochholdinger, 2006). Dabei wird von den beschriebenen Situationsanalysen, Handlungsabsichten und hypothetischen Handlungen auf Kompetenzen geschlossen. Situational Judgement Tests eignen sich somit insbesondere zur Erhebung von Handlungswissen bzw. prozeduralem Wissen anhand realitätsnaher Anforderungskontexte. Entsprechende Szenarien weisen darüber hinaus den Vorzug auf, dass sie die relevanten Leistungs- bzw. Kompetenzaspekte nicht isoliert, sondern in zusammenhängender Form und somit ökologisch valider abprüfen (Schaper, 2009).



## 6. Item- und Testentwicklung

Die Itemerstellung erfolgte in der Weise, dass zum einen alle Inhaltsbereiche des Rahmenmodells abgedeckt, zum anderen alle drei Anforderungsstufen des Faktors „Kognitive Aktivitäten“ bedient wurden. Im Rahmen der Test- bzw. Itementwicklung wurden in mehreren Expertenworkshops Aufgaben konstruiert mit dem Ziel, ein objektives, reliables und valides Messinstrument zu erstellen. Das Entwicklungsteam setzte sich dabei zusammen aus Experten mit unterschiedlichen Expertisefacetten: einer Erziehungswissenschaftlerin mit dem Schwerpunkt im Bereich Unterricht und Allgemeine Didaktik, einer Lehrerbildungsforscherin, einem Pädagogischen Psychologen und einem Psychologen mit dem Schwerpunkt diagnostische Forschungsmethoden. Alle an der Entwicklung des Instruments beteiligten Experten sind darüber hinaus in bedeutsamem Maße an der Ausbildung bzw. Lehre von zukünftigen Lehrkräften beteiligt.

Bei der Itementwicklung wurden die Aufgabenformate den kognitiven Anforderungen entsprechend gewählt. So wurden für die Stufe „Wissen reproduzieren, Verstehen“ eher Wissenstests verwendet, für die Stufe „Urteilen, Bewerten, Entscheiden“ eher Situational Judgement Tests, da diese besser geeignet sind, solche komplexen kognitiven Prozesse abzubilden. Pro Zelle wurden mehrere Aufgaben entwickelt, so dass eine Optimierung des Tests durch Itemselektion aufgrund entsprechender Itemkennwerte vorgenommen werden konnte. Sowohl die Aufgaben selbst als auch die Zuordnung zu den Zellen des Rahmenmodells wurden von Erziehungswissenschaftlern und Fachdidaktikern diskutiert und auf diese Weise inhaltlich validiert.

Im Folgenden werden in exemplarischer Form Aufgabenstellungen des Kompetenztests vorgestellt. Die Aufgabe EW06A (Abbildung 2) erfordert Verständnisleistungen in Bezug auf die Unterscheidung zwischen intrinsischer und extrinsischer Motivation. Das Item repräsentiert das Wissenstestformat und ist als Multiple-Choice-Aufgabe konstruiert. Inhaltlich wurde die Aufgabe zur Operationalisierung des Inhaltsbereichs „Unterricht und Allgemeine Didaktik“ mit dem Anforderungsbereich „Wissen reproduzieren, Verstehen“ des Rahmenmodells konzipiert.

EW	Man kann zwischen intrinsischer und extrinsischer Motivation unterscheiden.		
06	Bei welchem/n der folgenden Fälle handelt es sich um eine <i>intrinsische</i> Motivation?		
A	Ein Kind lernt vor einer Mathematikarbeit, weil es ...		
	A. für eine gute Note eine Belohnung erwartet		<input type="radio"/>
	B. einen Tadel für eine schlechte Note vermeiden möchte		<input type="radio"/>
	C. an mathematischen Problemen interessiert ist		<input type="radio"/>
	D. die sympathische Lehrerin nicht enttäuschen möchte		<input type="radio"/>

Abbildung 2: Aufgabe EW06A

Ein Beispiel für das Format des Situational Judgement Tests stellt die Aufgabe 20 dar (siehe Abbildung 3). Im ersten Teil (A) der Aufgabe wird ein praktisches Problem geschildert, welches mithilfe theoretischer Wissenskonzepte (Auswertungs- bzw. Interpretationsobjektivität) analysiert werden soll. Im zweiten Teil der Aufgabe (B) soll auf der Basis dieser Analyse ein Lösungsvorschlag generiert werden (Explikation und Präzisierung von Kriterien zur Bewertung der Klassenarbeiten). Die Aufgabestellung stellt eine Operationalisierung des Inhaltsbereichs „Schulentwicklung und Gesellschaft“ in Kombination mit dem Anforderungsbereich „Reflektieren, Anwenden“ dar.

- EW 20 Im Rahmen von Lernerfolgskontrollen sollen bestimmte Gütekriterien berücksichtigt werden. Stellen Sie sich bitte folgende Situation vor:  
*Eine Deutschlehrerin hat die Klassenarbeit einer zehnten Klasse bewertet. Der Klassendurchschnitt liegt bei 2,1. Ihrem Mann, der auch Deutschlehrer ist, kommt das sonderbar vor. Er bewertet dieselbe Arbeit und kommt auf einen Schnitt von 3,5*
- A In Bezug auf welches Gütekriterium liegt hier ein Problem vor?  
(Bitte antworten Sie in Stichpunkten)
- B Was kann die Lehrerin tun, damit andere Bewerter zu ähnlichen Ergebnissen wie sie kommen?  
(Bitte antworten Sie in Stichpunkten)

Abbildung 3: Aufgabe EW20

Im Folgenden sind drei weitere Aufgabenbeispiele aufgeführt (siehe Abbildung 4), die exemplarische Operationalisierungen für ausgewählte Felder der Inhalts- und Anforderungsmatrix des erziehungswissenschaftlichen Rahmenmodells wiedergeben. Die Beispiele, die zwei Wissenstestitems und ein Situational Judgement Test-Item umfassen, stehen für offene Antwortformate auf unterschiedlichen Anforderungsniveaus.

<b>Itemzuordnung</b>	„Schulentwicklung und Gesellschaft“ & „Wissen reproduzieren, verstehen“
<b>EW04</b>	Sie haben sich für den Lehrerberuf entschieden. Eine wichtige Aufgabe wird für Sie das Unterrichten sein. Welche weiteren Aufgabenbereiche kommen auf Sie als Lehrkraft zu? (Bitte antworten Sie in Stichpunkten)
<b>Itemzuordnung</b>	„Unterricht und Allgemeine Didaktik“ & „Reflektieren, Anwenden“
<b>EW24</b>	Nehmen Sie an, Sie sind ein Anhänger der kognitiven Lerntheorie. Wie würden Sie aus dieser Position argumentieren?  Man kann Schülerinnen und Schüler in ihren Lernprozessen unterstützen, indem ... (Antworten Sie in Stichpunkten!)
<b>Itemzuordnung</b>	„Erziehung und Bildung“ & „Urteilen, Bewerten, Entscheiden“
<b>EW10</b>	Stellen Sie sich bitte die folgende Unterrichtssituation vor:  <b>A</b> <i>Eine Deutschlehrerin hat sich vorgenommen, mit ihrer Klasse die Lektüre „Die Insel der blauen Delphine“ zu lesen. In der Unterrichtsstunde rebellieren ihre Schülerinnen und Schüler allerdings und wollen lieber ihre mitgebrachten Comics lesen. Um einen Kompromiss einzugehen, sagt die Lehrerin: „Na gut, ihr könnt erst eine Viertelstunde eure Comics lesen, dann lesen wir aber die Lektüre“.</i>  Was hat die Lehrerin aus lerntheoretischer Sichtweise falsch gemacht? (Bitte antworten Sie in Stichpunkten)

Abbildung 4: Aufgabenbeispiele für verschiedene Kombinationen von Inhalts- und Anforderungsbereichen

Auf dieser Grundlage konnten insgesamt 30 Aufgabenstellungen mit 111 Teil-Items für die Langversion des Tests zur Messung pädagogischer Kompetenzen entwickelt werden. Pro „Zelle“ der Matrix variiert die Anzahl der Items zwischen 3 bis 28. Auf der Basis von Pretests wurde die Bearbeitungsdauer des Tests auf 90 Minuten festgelegt. Außerdem wurden zwei Testformen (A und B) erstellt, um Reihenfolgeeffekte, Testlängeneffekte und ein Abschreiben bei der Erhebung zu vermeiden.

## 7. Testdurchführung

Um den Test zur Messung pädagogischer Kompetenzen hinsichtlich seiner psychometrischen Kennwerte zu erproben, wurde eine umfangreiche Piloterhebung in der Lehramtsausbildung an der Universität Paderborn durchgeführt. Die Erhebung erfolgte in zwei Wellen (Januar und Mai 2008) in jeweils drei unterschiedlichen Veranstaltungen des erziehungswissenschaftlichen Studiums (Grundstudium sowie Hauptstudium).

Die Stichprobe (N = 700) teilt sich auf in 514 weibliche (73,4 %) und 186 männliche Studierende (26,6 %). Tabelle 2 gibt die Verteilung der Testteilnehmer auf die verschiedenen Lehramtsstudiengänge wieder.

Tabelle 2: Verteilung der Studierenden auf die Studiengänge

<b>Grund-, Haupt-, Real- und Gesamtschule – Schwerpunkt Grundschule</b>	<b>Grund-, Haupt-, Real- und Gesamtschule – Schwerpunkt Haupt-, Real- und Gesamtschule</b>	<b>Gymnasium Gesamtschule</b>	<b>Berufskolleg</b>
182 (26 %)	189 (27 %)	275 (39.3 %)	49 (7 %)

Aus der Variable „Anzahl Fachsemester“ wurde für diese Pilotierung eine neue, binäre Variable codiert, die die Studierenden in Studienanfänger ( $\leq 3$ . Semester) und Studienfortgeschrittene ( $> 3$ . Semester) aufteilt, um zu überprüfen, ob es eine Progression bzw. Kompetenzentwicklung bezogen auf die einzelnen Anforderungsbereiche bereits innerhalb des Studiums gibt. Diese Frage ist bedeutsam, da die Wirksamkeit der Lehrerbildung bisher lediglich durch einen Vergleich zwischen Studium und Referendariat nachgewiesen wurde (vgl. Schmotz & Blömeke in diesem Band). Diese Aufteilung ist für die noch grundständig organisierte Lehrerbildung zweckmäßig, da sie der Struktur des erziehungswissenschaftlichen Studiums entspricht, bei dem das Grundstudium gemäß der Studienordnung der Erziehungswissenschaft in Paderborn nach dem dritten Semester abgeschlossen ist und das Hauptstudium mit dem 4. Semester beginnt. In der Stichprobe befinden sich 433 Studienanfänger (61,9 %) und 246 (35,1 %) Studienfortgeschrittene. Für Erhebungen im Rahmen von gestuften Studiengängen bietet sich eine Aufteilung in Bachelor und Master an, da sich diese beiden Phasen in struktureller und konzeptioneller Hinsicht unterscheiden.

Zur Auswertung des Tests wurde ein sog. Codebook erstellt, in dem die richtigen Antworten bzw. die passenden Antwortalternativen inklusive der jeweils zu vergebenden Punkte verzeichnet sind. Bei den offenen Antwortformaten wurden vom o. g. Expertenteam Musterlösungen entwickelt bzw. mögliche Antwortbeispiele mit Musterkodierungen versehen. Bei den offenen Fragen wurden dazu auch teilweise unterschiedliche Antwort- bzw. Qualitätsstufen definiert, für die jeweils Musterantworten und Punktwerte zur Bewertung vorgegeben wurden. Dies ist vorwiegend bei Aufgaben auf der Stufe „Urteilen, Bewerten, Entscheiden“ der Fall, bei denen von den Probanden die Analyse von Anforderungssituationen und die Generierung von Handlungsalternativen verlangt werden.

Ein Beispiel für Musterlösungen zu einer der oben vorgestellten Testaufgaben sah folgendermaßen aus (EW20; siehe Abbildung 2): EW20A: mangelnde Interpretationsobjektivität; EW20B: Erarbeiten und Transparentmachen einer Musterlösung bzw. eines Bewertungsschlüssels. Für jede richtige Antwort wurde bei dieser Aufgabe jeweils ein Punkt vergeben.

## 8. Ergebnisse

Zur psychometrischen Überprüfung des Instruments wurden zunächst die Reliabilität (Cronbachs  $\alpha$ ) sowie zusammenfassend die wichtigsten Itemkennwerte (Schwierigkeit, Streuung, Trennschärfe) berechnet. Zu einer ersten Bewertung der Validität des Tests wurden darüber hinaus Mittelwertsunterschiede zwischen Studierenden des Grundstudiums und des Hauptstudiums inferenzstatistisch überprüft (mittels t-Test). Die Annahme hierbei war, dass die Ergebnisse der Studierenden des Hauptstudiums signifikant besser sind als die Werte der Studierenden des Grundstudiums. Sofern der Test einen entsprechenden Kompetenzzuwachs durch Vergleich der beiden zeitlichen Kohorten ermitteln kann, wird dies als Validitätsnachweis für den Test gewertet.

Die Reliabilität für den gesamten Test beträgt  $\alpha = .868$ . In Tabelle 1 sind die Kennwerte für die Aufgaben, die oben als Beispiel angeführt sind, abgetragen:

Tabelle 3: Itemkennwerte der Beispielitems

Nr.	Item	Max	p	$S_D$	$r_{IT}$	$M_{GS}$	$M_{HS}$	t	sig.
	EW04	7	.34	0,186	.360	2,35	2,35	-,003	.998
	EW06AA	1	.63	0,483	.338	0,56	0,74	-4,915	.000
	EW06AB	1	.65	0,476	.343	0,58	0,77	-5,125	.000
	EW06AC	1	.61	0,489	.379	0,53	0,74	-5,651	.000
	EW06AD	1	.66	0,475	.220	0,57	0,79	-6,316	.000
	EW10A	1	.51	0,500	.147	0,55	0,46	2,254	.025
	EW20A	1	.40	0,491	.317	0,39	0,42	-,888	.375
	EW24	2	.06	0,174	.246	0,13	0,09	1,330	.184

Die anhand der Gesamtstichprobe berechneten Schwierigkeiten aller Items reichen von  $p = .04$  bis  $p = .96$ . Im Test sind somit sowohl sehr leichte als auch sehr schwierige Items vorhanden. Die mittlere Schwierigkeit liegt bei  $p = .47$ , also etwa im Bereich der Skalenmitte. Die Streuung der Items reicht von  $S_D = 0,15$  bis  $S_D = 0,50$ . Die mittlere Streuung beträgt  $S_D = 0,40$ . Die Trennschärfen befinden sich im Bereich von  $r_{IT} = .019$  bis  $r_{IT} = .467$ , wobei zwei Items auch negative Trennschärfen aufweisen.

Bei insgesamt 40 Items konnte festgestellt werden, dass Studierende des Hauptstudiums signifikant bessere Ergebnisse erzielen als Studierende des Grundstudiums. Allerdings sind bei 17 Items Studierende des Grundstudiums besser als Studierende des Hauptstudiums. Damit kann die Validität des Testinstruments im Sinne der Feststellung von Kompetenzzuwächsen im Verlauf des Studiums nur eingeschränkt bestätigt werden.

## 9. Entwicklung einer Kurzversion und erste Ergebnisse

In einem weiteren Schritt wurde eine Kürzung des Instruments vorgenommen mit dem Ziel, die pädagogische Kompetenz als Teil eines umfassenderen Kompetenzprofils von Lehrkräften zusammen mit fachlichen und fachdidaktischen Kompetenzen zu erfassen. Weiterhin sollen hierbei neben den kognitiven Teilaspekten (fachliches, fachdidaktisches und erziehungswissenschaftliches Wissen) auch nicht-kognitive Aspekte wie Werthaltungen und motivationale Orientierungen als Elemente professioneller Handlungskompetenz von Lehrkräften berücksichtigt bzw. erfasst werden.

Um auf der Basis der Langversion eine Kurzversion des Instruments zu erstellen, wurde eine Itemselektion unter Berücksichtigung folgender Kriterien vorgenommen:

- Elimination von Items, die große inhaltliche Ähnlichkeiten zu anderen Items aufweisen (dieser Selektionsschritt erfolgte in Zusammenarbeit mit zwei Experten aus der Erziehungswissenschaft)
- Elimination von Items, die einen spezifischen Standortbezug aufweisen (Wissensinhalte, die in erster Linie im Lehramtsstudium der Universität Paderborn vermittelt werden)
- Berücksichtigung itemanalytischer Kennwerte (Schwierigkeit, Streuung, Trennschärfe) bei der Selektion.

Die resultierende Kurzversion des Tests besteht aus 11 Aufgabenstellungen mit insgesamt 30 Teilitems. Exemplarische Durchführungen mit der Kurzversion zeigten, dass durch die Itemselektion eine durchschnittliche Bearbeitungszeit von 30 Minuten erreicht wird.

Die Reliabilitäts- und Itemkennwertberechnungen zur psychometrischen Überprüfung der Kurzform wurden an derselben Stichprobe wie die der Langform vorgenommen. Die Reliabilität der Kurzversion beträgt  $\alpha = .708$  und ist damit noch als zufrieden stellend zu bewerten. Weiterhin wurde überprüft, ob dieser Wert allein durch die Itemreduktion der Skala resultiert. Setzt man den Faktor der Verkürzung in die Spearman-Brown-Formel ein (Bühner, 2006, S. 139), so errechnet sich eine Reliabilität von .647. Da die empirisch ermittelte Reliabilität der Kurzversion höher ausfällt als die mithilfe der Verkürzungsformel berechnete Reliabilität, kann gefolgert werden, dass das Instrument durch die Itemselektion weiter optimiert werden konnte.

Die Schwierigkeitskennwerte der Kurzform reichen von  $p = .06$  bis  $p = .66$ , wobei die mittlere Schwierigkeit bei  $p = .371$  liegt. Somit ist die Kurzform etwas leichter als die Langform. Gleichzeitig konnten aber auch extrem „schwierige“ Items eliminiert werden. Die Streuungen der verbleibenden Items der Kurzver-

sion liegen zwischen  $S_D = 0,17$  und  $S_D = 0,50$ . Die mittlere Streuung fällt gegenüber der Langform mit  $S_D = 0,410$  etwas höher aus.

Die Trennschärfen der Items liegen schließlich zwischen  $r_{IT} = .043$  und  $r_{IT} = .437$ . Durch die Verfahrenskürzung konnten zwar Items mit negativen Trennschärfen eliminiert werden. Allerdings besteht in Bezug auf die Itemtrennschärfen weiter Optimierungsbedarf.

In Bezug auf die Überprüfung, inwieweit mithilfe des Instruments Kompetenzunterschiede bzw. -zuwächse zwischen Grund- und Hauptstudium ermittelt werden können, ist festzustellen, dass nach der Itemreduktion bei 14 Items die Studierenden des Hauptstudiums signifikant besser sind als die Studierenden des Grundstudiums; nur noch bei zwei Items ist der Effekt entgegengesetzt.

## 10. Diskussion und Ausblick

Zielsetzung der Studie war die Entwicklung eines Messinstruments zur Erfassung pädagogischer Kompetenzen, welches in der Lage ist, den Kompetenzerwerb bezüglich des erziehungswissenschaftlichen Teils des Studiums abzubilden und das als Evaluationsinstrument im Rahmen dieses (Teil-)Studiengangs eingesetzt werden kann. Die vorgestellten Ergebnisse verdeutlichen, dass es insgesamt gelungen ist, ein in zufrieden stellendem Maße objektives und reliables Messinstrument zur Erfassung pädagogischer Kompetenzen im Lehramtsstudium der Universität Paderborn zu entwickeln. Einschränkend muss allerdings zum gegenwärtigen Zeitpunkt auch festgestellt werden, dass der untersuchte Validitätsnachweis bei der Langversion des Instruments noch nicht in ausreichender Form erbracht worden ist (bei 17 Items sind Studierende des Grundstudiums besser als Studierende des Hauptstudiums). Überprüft werden muss, inwieweit dies tatsächlich an der mangelnden Validität des Instruments liegt. Das Ergebnis könnte auch dadurch bedingt sein, dass der Test Wissensinhalte abfragt, die primär im Grundstudium behandelt werden und dann nicht im Sinne eines aufbauenden Spiralcurriculums ausreichend wiederholt bzw. konsolidiert werden. Damit wird möglicherweise bereits auf Schwachstellen des Kompetenzerwerbs im erziehungswissenschaftlichen Studium hingewiesen, die gegebenenfalls auch mit der Modularisierung des Studiums und mit einem geringen Gesamtstudienvolumen zusammen hängen könnten. Der Erziehungswissenschaft steht im Vergleich zu den Fächern ein geringeres Studienvolumen zur Verfügung, das – gemäß den Vorgaben – modularisiert wird und damit in Inhaltszusammenhänge von 6-10 SWS gruppiert wird. Damit findet der Kompetenzerwerb bezogen auf einen Inhaltsbereich kompakt in einem 1-2-semesterigen Zusammenhang statt. Für ein Vertiefungsmodul zu einem späteren Zeitpunkt zu demselben Inhaltsbereich verbleiben angesichts der zu erwerbenden Breite an Kompetenzen keine zeitlichen Ressourcen. Eine weitere Erklärung für das bessere Abschneiden der Studierenden im Grundstudium

könnte auch in der Art der Lerngelegenheiten und -methoden liegen, die gegebenenfalls nicht dazu angetan sind, den Wissenserwerb und damit den Kompetenzzuwachs nachhaltig zu sichern. Hier bedarf es somit weiterer Analysen und Erhebungen zur Klärung der Verläufe bei der Aneignung und Konsolidierung von spezifischen Wissensinhalten und Kompetenzfacetten im erziehungswissenschaftlichen Studium.

Ein weiterer Arbeitsschritt zur Validierung des Instruments stellt die Überprüfung der dem Instrument zugrunde liegenden Strukturierung pädagogischer Kompetenzen dar. Hierbei soll überprüft werden, ob sich die im Rahmenmodell angenommene Strukturierung pädagogischer Kompetenzen auch empirisch mithilfe konfirmatorischer Faktorenanalysen bestätigen lässt. Das entsprechende Modell geht aus von einer Inhaltsdimension mit drei moderat korrelierten Faktoren sowie einer Anforderungsdimension mit drei Faktoren, die aufeinander aufbauen.

Ein weiterer bedeutsamer Schritt stellt darüber hinaus die kriteriumsbezogene Validierung des Tests dar. Da mithilfe des Verfahrens *professionelle* Kompetenzen erhoben bzw. diagnostiziert werden sollen, gilt es zu überprüfen, inwieweit unterschiedliche Ausprägungen in Testleistungen auch mit entsprechenden Ausprägungen im Unterrichtshandeln korrespondieren bzw. korrelieren. Entsprechende Kriteriumsleistungen könnten mithilfe von videobasierten Aufzeichnungen und Bewertungen des Lehrhandelns im Unterricht bei Lehramtsstudierenden im Praktikum bzw. bei Referendaren erhoben werden.

Die Kurzform weist zufriedenstellende Kennwerte in Bezug auf Schwierigkeit und Reliabilität des Instruments auf und kann damit vorbehaltlich weiterer Validitätsprüfungen ebenfalls zur Messung pädagogischer Kompetenzen im Lehramtsstudium verwendet werden. Durch die erhebliche Reduktion der Bearbeitungszeit auf eine halbe Stunde ist dieses Verfahren zudem geeignet, im Zusammenhang mit weiteren Teilinstrumenten das Profil professioneller Kompetenz bei angehenden Lehrkräften umfassender zu messen. Mithilfe solcher Untersuchungsansätze kann analysiert werden, welche Beziehungen die pädagogische Kompetenz zu den anderen Facetten professioneller Handlungskompetenz, d. h. Fachwissen und fachdidaktisches Wissen, sowie zu professionellen Werthaltungen und motivationalen Orientierungen von Lehrkräften aufweist.

Eine weitere Fragestellung betrifft die Konstruktvalidität des Instruments: Inwieweit ist das Verfahren in der Lage, nicht nur am Standort Paderborn, sondern auch an anderen Ausbildungsstandorten die Qualität der erziehungswissenschaftlichen Ausbildung valide zu erfassen und zu evaluieren? Ist das Instrument darüber hinaus vergleichbar mit anderen Instrumentarien zur Messung pädagogischer Kompetenzen? Erste Antworten hierzu soll die sog. LEK-Studie (LEK: *Längsschnittliche Erhebung allgemeiner pädagogischer Kompetenzen* von Lehramtsstudieren-



den) erbringen, in der die Kurzform parallel zum Instrumentarium von König und Blömeke (im Druck) an verschiedenen Universitätsstandorten eingesetzt wird. Die Studie ist zudem als Längsschnitt geplant, sodass aussagekräftigere Ergebnisse zum Entwicklungsverlauf beim Erwerb pädagogischer Kompetenzen im Lehramtsstudium erhoben werden.

### Literatur

- Abs, H. J., Döbrich, P., Vögele, E. & Klieme, E. (2005). Skalen zur Qualität der Lehrerbildung – Dokumentation der Erhebungsinstrumente: Pädagogische Entwicklungsbilanzen an Studienseminaren (PEB-Sem). Frankfurt am Main: DIPF.
- Baer, M., Dörr, G., Fraefel, U., Kocher, M., Küster, O., Larcher, S., Müller, P., Sempert, W. & Wyss C. (2007). Werden angehende Lehrpersonen durch das Studium kompetenter? Unterrichtswissenschaft, 35, 15-47.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. Zeitschrift für Erziehungswissenschaft, 9, 469-520.
- Baumert, J., Kunter, M., Brunner, M., Krauss, St., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In M. Prenzel et al. (Hrsg.), PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs (S. 314-354). Münster: Waxmann.
- Blömeke, S. (2004). Empirische Befunde zur Wirksamkeit der Lehrerausbildung. In S. Blömeke, P. Reinhold, G. Tulodziecki & J. Wildt (Hrsg.), Handbuch Lehrerbildung (S. 59-91). Bad Heilbrunn/Braunschweig: Klinkhardt.
- Blömeke, S., Felbrich, A. & Müller, C. (2008a). Messung erziehungswissenschaftlichen Wissens angehender Lehrkräfte. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematik-Studierender und -referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung (S. 171-193). Münster: Waxmann.
- Blömeke, S., Müller, C. & Felbrich, A. (2008b). Entwicklung des erziehungswissenschaftlichen Wissens und der professionellen Überzeugungen in der Lehrerausbildung. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematik-Studierender und -referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung (S. 301-326). Münster: Waxmann.
- Blömeke, S. Reinhold, P., Tulodziecki, G. & Wildt, J. (Hrsg.). (2004). Handbuch Lehrerbildung. Bad Heilbrunn/Braunschweig: Klinkhardt.

- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 177-212). Göttingen: Hogrefe.
- Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Dubberke, T., Jordan, A., Klusmann, U., Tsai, Y.-M. & Neubrand, M. (2006). Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? *Zeitschrift für Erziehungswissenschaft*, 9 (4), 521-544.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Darling-Hammond, L. & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world. What teachers should learn and be able to do*. San Francisco: Jossey-Bass.
- Dubberke, T., Kunter, M. & McElvany, N. (2008). Pedagogical knowledge as an aspect of teacher competence: Conceptualization and test construction. *International Journal of Psychology*, 43 (3/4), 68.
- Frey, A. (2008). Kompetenzstrukturen von Studierenden in der ersten und zweiten Phase der Lehrerbildung. Eine nationale und internationale Standortbestimmung. Landau: Empirische Pädagogik.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen*. Münster: Waxmann.
- Helmke, A. (2003). *Unterrichtsqualität erfassen, bewerten, verbessern*. Seelze: Kallmeyersche Verlagsbuchhandlung.
- Hilligus, A. H., Reinhold, P., Rinkens, H.-D. & Tulodziecki, G. (2004). *SPEE: Standards – Profile – Entwicklung – Evaluation: Innovation in der Lehrerbildung an der Universität Paderborn*. Paderborn: PLAZ, PLAZ-Forum: Lehrerbildung und Schule in der Diskussion, Heft Nr. 5.
- Hilligus, A. H., Rinkens, H.-D. (Hrsg.). (2006). *Standards und Kompetenzen – neue Qualität in der Lehrerbildung?* Berlin: LIT.
- KMK (2004). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Standards für die Lehrerbildung: Bildungswissenschaften*. Beschluss der Kultusministerkonferenz vom 16.12.2004.
- KMK (2008). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung*. Beschluss der Kultusministerkonferenz vom 16.10.2008.
- König, J. & Blömeke, S. (im Druck). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerbildung. *Zeitschrift für Erziehungswissenschaft*.

- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2003). Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Berlin: BMBF.
- Nerdinger, F., Blickle, G. & Schaper, N. (2008). Lehrbuch Arbeits- und Organisationspsychologie. Heidelberg: Springer.
- Oser, F. & Oelkers, J. (Hrsg.). (2001). Die Wirksamkeit der Lehrerbildungssysteme. Chur: Rüegger.
- Schaefer, C. (2002). Forschung zur Lehrerausbildung in Deutschland – eine bilanzierende Übersicht der neueren empirischen Studien. Schweizerische Zeitschrift für Bildungswissenschaften, 24, 65-88.
- Schaper, N. (2006). (Arbeits-)Psychologische Konzepte zur Modellierung und Erfassung von Kompetenzen in der Lehrerausbildung. Vortrag zum Workshop der Universität Hannover und der Hochschulrektorenkonferenz „Kompetenzen und Kompetenzentwicklung in der Lehrerbildung“. Hannover, 31.03.2006. Verfügbar unter: [http://www.hrk.de/bologna/de/download/dateien/vortrag\\_schaper\\_kompetenzen310306.pdf](http://www.hrk.de/bologna/de/download/dateien/vortrag_schaper_kompetenzen310306.pdf) [18.12.2006]
- Schaper, N. (2009). (Arbeits-)Psychologische Kompetenzforschung. In M. Fischer & G. Spöttl (Hrsg.), Forschungsperspektiven in Facharbeit und Berufsbildung (S. 91-115). Frankfurt: Peter Lang.
- Schubarth, W. & Polenz, P. (Hrsg.) (2006). Qualitätsentwicklung und Evaluation in der Lehrerbildung. Die zweite Phase: das Referendariat. Potsdam: Universitätsverlag.
- Shulman, L. S. (1985). Paradigms and research programs in the study of teaching. A contemporary perspective. In M. C. Wittrock (Ed.), Handbook of Research on Teaching, 3. Auflage, (S. 3-36). New York: Macmillan.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15, 4-14.
- Terhart, E. (2002). Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz. Münster, ZKL-Texte Nr. 24.
- Tulodziecki, G. (2004). Bildungsstandards im erziehungswissenschaftlichen Studium für Lehrämter. Paderborn: PLAZ, PLAZ-Forum: Lehrerausbildung und Schule in der Diskussion, Heft Nr. 5.
- Ulbricht, T., Schaper, N. & Hochholdinger, S. (2007). Entwicklung von Instrumenten zur Evaluation einer kompetenzorientierten Lehrerausbildung. In D. Lemmermöhle, u. a. (Hrsg.), Professionell lehren – erfolgreich lernen (S. 170-182). Münster u. a.: Waxmann.
- Weekley, J. A. & Ployhart, R. E. (2006). Situational judgement tests. Theory, measurement and application. Mahwah, N.J.: Lawrence Erlbaum.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17-31). Weinheim: Beltz.

*Anschrift der Autoren:*

Prof. Dr. Niclas Schaper, Lehrstuhl für Arbeits- und Organisationspsychologie, Fakultät für Kulturwissenschaften, Universität Paderborn, Warburger Straße 100, 33098 Paderborn, Tel.: 05251-602901, Fax: 05251-60352,  
E-Mail: niclas.schaper@uni-paderborn.de

Dr. Annegret Helen Hilligus, Universität Paderborn, Zentrum für Bildungsforschung und Lehrerbildung (PLAZ), Peter-Hille-Weg 42, 33098 Paderborn,  
E-Mail: plaz-hi@upb.de (Hilligus)

Dr. Andreas Seifert, Universität Paderborn, Zentrum für Bildungsforschung und Lehrerbildung (PLAZ), Peter-Hille-Weg 42, 33098 Paderborn,  
E-Mail: seifert@plaz.upb.de (Seifert)